

Promoting Acceleration of Comprehension and Content Through Text in High School Social Studies Classes

Jeanne Wanzek

Florida State University, Tallahassee, Florida, USA

Elizabeth A. Swanson, Greg Roberts, and Sharon Vaughn

The University of Texas at Austin, Austin, Texas, USA

Shawn C. Kent

Florida State University, Tallahassee, Florida, USA

Abstract: The purpose of this study was to evaluate the efficacy of Promoting Acceleration of Comprehension and Content Through Text intervention implemented with 11th-grade students enrolled in U.S. History classes. Using a within-teacher randomized design, the study was conducted in 41 classes (23 treatment classes) with 14 teachers providing the treatment. Students in the treatment condition performed significantly better than students in the typical instruction comparison condition on a measure of content acquisition at posttest ($ES = 0.36$), as well as 4 ($ES = .22$) and 12 ($ES = .24$) weeks following treatment. There were no differences between treatment and comparison groups on measures of social studies reading comprehension or more general reading comprehension.

Keywords: Team-base learning, content area literacy, secondary instruction

High school students who are prepared to successfully complete college courses and/or enter the workforce with greater job choices and higher income potential possess a variety of academic skills including independent learning, a command of content knowledge, reading with comprehension, and using text evidence to support their writing and speaking (Hein, Smerdon, Lebow, & Agus, 2012; New York City Department of Education, 2013). The social studies offer one discipline for students to acquire these skills. The goal of social studies instruction is to prepare knowledgeable and informed citizens, and its content provides rich, complex concepts and events, often with divergent points of view, that require the ability to read and think critically for understanding (Barton & Levstik, 2003; Lee & Spratley, 2010; National Council for the Social Studies, 2008; Paxton, 1999). To assist teachers and students in accomplishing this goal, the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) complement social studies content by providing anchors for infusing the key literacy skills needed to successfully build knowledge in history and social studies. Within the Common Core State Standards at the high school level, social studies students are taught to (a) determine and summarize central historical events and ideas; (b) provide explanations for these ideas citing text evidence; (c) learn domain-specific vocabulary and analyze how word choice refines the intent of authors' documents; and (d) evaluate differing

points of view among key players in history by assessing author claims, reasoning, and evidence.

Infusing social studies instruction with literacy components is a lofty requirement for many social studies teachers, who often view their primary role as a content delivery expert and may be less comfortable with their proficiency in teaching students how to improve their reading comprehension (Hall, 2005; O'Brien, Moje, & Stewart, 2001; O'Brien, Stewart, & Moje, 1995). A recent observation study (Swanson et al., in press) of middle and high school social studies classes provides evidence that less than 10% of observed classroom time was allocated to text reading in high school social studies and was lowest of all in 11th grade, where text reading consumed less than 5% of class time. Other features of literacy instruction for understanding text in the social studies discipline were observed to a very limited extent. For example, vocabulary instruction rarely extended beyond presenting definitions, and most comprehension instruction entailed accessing prior learning and teacher questioning of understanding.

Students are unlikely to develop sophisticated content knowledge without intentional instruction in comprehending and integrating text information in the content area (Berkeley, Marshak, Mastropieri, & Scruggs, 2011; Lee & Spratley, 2010; Stahl, Hynd, Glynn, & Carr, 1996). Admittedly, designing an intervention that maintains a social studies teacher's role as content expert while providing evidence-based instructional approaches to teach students vocabulary and comprehension in the domain is a challenging prospect (Heller & Greenleaf, 2007; Shanahan & Shanahan, 2008; Vaughn & Fletcher, 2012). One recently developed intervention, Promoting Acceleration of Comprehension and Content Through Text (PACT), is designed to facilitate this process for middle and secondary social studies teachers aiming to engage students in the content through vocabulary, text reading, and discussion-based processes in order to improve student understanding of the content. The intervention was designed in an iterative design experiment process with 12 middle and high school social studies teachers (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). In the design experiments, teacher and student use of text and discussion-based elements within the existing social studies curricula were examined, and implementation was refined with each iteration (Vaughn et al., 2013). Following these design experiments a pilot study of the intervention was conducted to examine the use and feasibility within an instructional unit. Observation data, teacher and student interviews, and student work were collected to analyze the implementation and further refine the intervention. From this preliminary work, the PACT intervention was devised to be embedded in teachers' content instruction aligned with existing district and state standards (Vaughn et al., 2013).

The PACT intervention consists of five components: (a) a comprehension canopy to introduce the unit, activate background knowledge, and provide an overarching unit topic question to be addressed through the upcoming content; (b) instruction and continual review of essential words and concepts for the unit content; (c) knowledge acquisition through critical reading of primary and secondary text sources; (d) team-based learning (TBL; Michaelsen & Sweet, 2001) comprehension checks to incorporate collaborative discussions among students regarding the new content, checks on comprehension of the critical content, and reteaching of content as needed; and (e) TBL knowledge application to provide opportunities for students to apply the content to problem-solving or perspective-taking tasks.

In a recent series of studies (see Vaughn et al., 2013; Vaughn et al., in press), authors examined the efficacy of the PACT intervention related to content knowledge and reading comprehension outcomes for middle school students. In the first study, Vaughn et al.

(2013) worked with five eighth-grade social studies teachers to implement the PACT intervention for three instructional units. In this experimental study, 27 class sections were randomly assigned within teacher to either PACT or a comparison group of the teachers' typical instruction. Following the unit instruction, the eighth graders assigned to the PACT condition outperformed students assigned to the comparison condition on measures of content acquisition ($ES = 0.17$), comprehension of text related to social studies ($ES = .29$), and the standardized reading comprehension ($ES = 0.20$). In a replication study the following year (Vaughn et al., in press), authors reported confirmation of the impact of the PACT intervention on content acquisition ($ES = 0.32$) that was also sustained at 4- and 12-week follow-up testing ($ES = 0.29$ and 0.26 , respectively) for students in eighth-grade social studies. However, the reading comprehension findings were not replicated, and no differences between groups were noted on comprehension of social studies text or general reading comprehension.

A related study (Wanzek et al., 2014) examined the efficacy of the TBL component at the 11th-grade level. Twenty-six classes were randomly assigned within teacher to receive TBL or typical practice instruction. The TBL classes embedded practices to support students in engaging in discourse about course content, application of content to solve problems, and the use of evidence to support responses. After three instructional units, findings indicated significant differences in favor of the TBL group on content acquisition ($ES = 0.19$). There were also differences in student response to the treatment. Students who were classified with high or moderate pretest scores benefitted from the treatment, whereas a group of students classified with low pretest scores did not benefit from the TBL treatment.

Thus, the PACT intervention has been examined at the middle school level with replicated findings for improved content acquisition. In addition, the TBL component of PACT has been noted to improve average student content acquisition at the high school level. However, the PACT intervention has not been examined in high school classes. The purpose of this study was to investigate the efficacy of the PACT intervention among 11th-grade students in social studies classes. We hypothesized that (a) students in the PACT treatment would outperform the students in the comparison condition on a measure of content acquisition; (b) comprehension of text related to the social studies domain would be enhanced for students receiving the PACT treatment; and (c) scores on a more distal, general reading comprehension measure would not differ between the treatment and comparison groups.

METHOD

Study Design

An experimental study was implemented with randomization of social studies classes to treatment or comparison, blocking on teacher. This design systematically controls for teacher effects and is aligned with the previous studies of PACT (e.g., Vaughn et al., 2013). Each teacher had at least one treatment and one comparison class. When a teacher had an odd number of class sections, the additional class was assigned to the treatment condition (e.g., five sections = three treatment sections and two comparison sections). There were 23 classes randomly assigned to treatment and 18 classes randomly assigned to comparison.

Table 1. Student demographics by condition

Demographic	Treatment ^a		Comparison ^b	
	N	%	N	%
Gender				
Male	253	54.1	158	45.5
Female	198	42.3	180	51.9
Missing	17	3.6	9	2.6
Race and ethnicity				
Hispanic	116	24.8	66	19.0
Alaskan Native/ Native American	48	10.3	27	7.8
Asian	4	.9	4	1.2
Black	106	22.6	91	26.2
Pacific Islander	1	.2	0	0
White	262	56.0	195	56.2
Multiracial	24	5.1	15	4.3
Missing	23	4.9	15	4.3
FRL status				
Yes	217	48.3	143	42.6
No	232	49.6	193	55.6
Missing	19	4.1	11	3.2
Disability status				
Yes	39	8.3	35	10.1
No	408	87.2	302	87.0
Missing	21	4.5	10	2.9
LEP				
Yes	12	2.6	9	2.6
No	437	93.4	327	94.2
Missing	19	5.5	11	3.2

Note. FRL = free or reduced lunch; LEP = limited English proficiency.

^a*n* = 468. ^b*n* = 347.

Participants

All of the U.S. History teachers in five high schools from four school districts located in the Southwest (Site 1) and Southeast (Site 2) participated in the study. Thus, a total of 14 teachers and their 41 United States History classes were included.

Teachers. Nine male and five female teachers participated. All of the teachers held a bachelor's degree with certification. Three of the teachers also held a master's degree. Teachers' ethnicity was White (*n* = 11) or Black (*n* = 3). The range of teaching experience prior to the study year was 2 to 27 years (*M* = 11.36 years, *SD* = 7.64), with a mean of 8.64 years teaching 11th grade.

Students. Table 1 provides the demographic information for the 815 students (411 male students) who consented to participate in the study. This sample consisted of 22% Hispanic race and 78% not Hispanic. Ethnicity was largely White (56%), Black (24%), and Alaskan Native/Native American (9%). Forty-four percent of these students qualified for free or reduced lunch, and 9% qualified for special education services. There were significantly

more male participants in the treatment group. There were no other significant differences between study conditions on any of the demographic variables.

Intervention Procedures

Teachers implemented three instructional units (Gilded Age, Imperialism and World War I, The Twenties) in their treatment and comparison classes. Each unit lasted 10 instructional days, for a total of thirty 45- to 50-min classes of implementation. Teachers used their typical practices during the comparison classes so that students received the instructional units as they would typically occur. In the treatment classes, the same content was taught, but teachers implemented the PACT components during the instruction. Thus, the treatment and comparison classes received the same content coverage aligned with the state and district standards over the same period. The contrast of interest was the delivery of the content, PACT, or typical practice, and not the content or curriculum provided.

Teachers attended a 1-day professional development workshop prior to the start of the study. The workshop covered PACT implementation as well as study design features related to maintaining the treatment only in the treatment classes. Each of the intervention components were described, modeled, and practiced during the day, including procedures for facilitating student discussion, critical reading, and use of text evidence to support claims.

During implementation of the units, teachers were provided in-class coaching support in their treatment classes. Support included assistance with planning units, modeling components, co-teaching, and observation and feedback on implementation. The coaching support was based on need but was provided at minimum once per week for each teacher.

Description of the Treatment

The PACT social studies treatment consisted of five components: (a) comprehension canopy, (b) essential words and concepts, (c) critical reading of primary and secondary sources, (d) TBL comprehension checks, and (e) TBL knowledge application. Teachers were provided with a unit schedule for implementing the PACT components within their social studies content instruction. All materials needed to implement the PACT components were provided to the teachers.

Comprehension Canopy. The comprehension canopy was designed to engage students in the topic, activate prior knowledge, and provide purpose for learning throughout the unit. A three-part unit introductory routine was followed on the 1st day of each unit. First, teachers introduced the unit with a brief discussion of questions related to prior units and designed to connect the upcoming unit topic to prior learning. Second, a unit introduction springboard activity such as a short video, photo, or map related to the unit was introduced with a series of questions designed to build background knowledge and connect the topic to students' experiences and prior knowledge. Third, an overarching question to guide the unit activities was introduced (e.g., During the Gilded Age, how did the economic, political, and social landscape of America change?). Throughout the unit, the unit question was reviewed continually to allow time for students to address the question as they learned the material develop a complete answer by the end of the unit.

Essential Words and Concepts. On the 1st day of each unit, four or five vocabulary words and concepts were introduced to students. The words were selected based on their utility for learning the content, frequency in the social studies and across other subject areas, and the way in which they address higher level concepts within the unit (e.g., industrialization, prosperity, social revolution). Each essential word was introduced with a student-friendly definition, visual representation, related words, sentences as examples of the word use, and question prompts to guide a brief discussion of the word meaning in context (e.g., What is one thing you would do for your family if you had greater prosperity? How is a social revolution different from a political revolution?). Throughout the unit, the essential words were reviewed during beginning of class warm-up activities that required students to use the word or word meaning to address unit related questions. In addition, the essential words were integrated in the critical readings, TBL comprehension checks, and TBL knowledge application in order to provide students with multiple exposures and practice opportunities for using the words.

Critical Reading. On Days 3, 5, and 7, teachers facilitated critical readings of primary and secondary text sources of the content. Whole-class, small-group, paired, or individual silent reading was implemented with stopping point questions designed to check for understanding, allow for student discussion, and provide space in the text for students to take notes on text meaning and connections to the comprehension canopy, essential words, and content learned.

TBL Comprehension Checks. Comprehension checks, including individual and team checks of content understanding, were implemented on Days 4 and 8 of the unit. Quiz questions were designed to require students to integrate and evaluation content learned, with the goal of eliciting discussion during the team checks. Students began by taking the quiz individually, providing teachers with a measure of student understanding and retention of the content taught.

Students then took the quiz again in heterogeneous teams of three to four students that were strategically formed by teachers prior to implementation of the study units. Students remained in the same teams throughout all three units. During the team quiz, students discussed each question, identified evidence from their course materials to support their answers, and determined a team answer for the question. Immediate feedback for answers was provided through the use of a scratch-off card keyed to the comprehension check answers, displaying a star for the correct answer. If a team selected the incorrect answer, the members would discuss the question and evidence further until they arrived at the correct answer. Teachers facilitated discussion and use of evidence during the team comprehension checks. Teams were also able to challenge a question or answer if they could provide evidence for one of their answers that was marked as incorrect. The teacher then discussed the evidence with the team and determined whether to accept the answer. Teachers used the data from the individual and team checks to identify misunderstandings and content in need of further clarification or reteaching.

TBL knowledge Application. At the end of the unit, students worked in their teams to complete a knowledge application activity requiring them to discuss the content to address a problem or make a specific choice (e.g., Imagine you are an advisory committee to a Gilded Age president. As a team, make a recommendation regarding whether or not the United States should limit immigration. Provide at least two economic, two social and two political reasons in support of your recommendation.). Students were asked to identify

evidence from the texts and notes to support their points, record key information from the discussions on team planning sheets, and report team decisions/rationale to their peers.

Observation of Fidelity of Treatment and Comparison Classes

We collected data on instruction in the treatment and comparison classes. These data allowed us to examine the extent to which the treatment was implemented in the treatment classes as well as the extent to which there were any PACT elements implemented in the comparison classes. As a result, we can determine the achieved strength of the treatment implementation and its relationship to the intended treatment (Hulleman & Cordray, 2009).

To capture implementation of the instruction, one treatment and one comparison class were randomly selected for audio-recording. We audio-recorded every day of instruction (30 days) in the selected classes for each of the teachers. For the purposes of determining fidelity of implementation (treatment condition) and the extent to which treatment elements occurred in the comparison condition, two recordings of each intervention component were randomly selected in each treatment class for a total of 10 treatment audiorecordings per teacher. Ten consecutive audiorecordings were also selected for each comparison class.

Using a fidelity instrument of the five components of the treatment, each audiorecording was coded by a research team member who was blind to condition. Coders rated the instruction on a 4-point Likert-type scale to indicate the extent to which each PACT component was implemented (e.g., 1 = *component not observed*, 2 = *few elements and procedures completed*, 3 = *a majority of the elements and procedures completed*, 4 = *all of the required elements and procedures completed*).

Interrater reliability on the fidelity tool was established using a gold standard method (Gwet, 2001). Six coders were trained on the PACT intervention and independently coded an audiorecording from a previous study. Codes were compared with a gold standard that was established on the same audio-recording by two senior researchers. Coders were required to establish reliability of 90% or higher. Table 2 provides the implementation data for the treatment and comparison classes.

Treatment Classes. The treatment was implemented at a mean rating of average with a majority of the required elements and procedures implemented. Mean implementation in the treatment classes across components was 2.91. Ratings of the implementation of the comprehension canopy, essential words, warm-ups to review essential words, and TBL comprehension checks all averaged 3.00 or above. Implementation of critical reading of text ($M = 2.37$) and TBL knowledge application was generally more difficult for teachers ($M = 2.52$).

Comparison Classes. A total of 117 classroom observations across the three units indicated very little overlap between teachers' typical practices and the PACT intervention. The average implementation was 1.05 (or *not observed*) in the comparison observations compared to 2.91 (or the *majority of components implemented*) in the treatment classes, providing the relative strength of the PACT intervention in the treatment classes. Some teachers did have a practice similar to the warm-up where they gave students an activity related to the content to work on for the first few minutes of the class period and discussed responses. The activities were generally not related to review and application of vocabulary concepts as they were in the PACT intervention. Thus, the large majority of PACT elements

Table 2. Frequency of fidelity observations in treatment and comparison classrooms

Implementation	CC		EW		WU		CRT		TBLCC		TBLKA	
	<i>n</i> = 28	%	<i>n</i> = 28	%	<i>n</i> = 30	%	<i>n</i> = 28	%	<i>n</i> = 29	%	<i>n</i> = 29	%
4 = High	8	28.57	12	42.86	18	60.0	1	3.57	7	24.14	1	3.45
3 = Average	13	46.43	12	42.86	4	13.33	12	42.86	20	68.97	14	48.28
2 = Low	6	21.43	4	14.29	2	6.66	10	35.71	1	3.45	13	44.83
1 = Not observed	1	3.57	0	0	6	20.0	4	14.29	1	3.45	1	3.45
Treatment Classrooms												
Comparison Classrooms												
4 = High	<i>n</i> (117)	%	<i>n</i> (117)	%	<i>n</i> (117)	%	<i>n</i> (117)	%	<i>n</i> (117)	%	<i>n</i> (117)	%
3 = Average	0	0	0	0	3	2.56	0	0	0	0	0	0
2 = Low	0	0	1	.85	1	.85	0	0	0	0	0	0
1 = Not observed	1	.85	2	1.71	10	8.55	4	3.42	0	0	2	1.71
	116	99.15	114	97.40	103	88.03	113	96.6	117	100	115	98.29

Note. CC = comprehension canopy; EW = essential words; WU = warm-up; CRT = critical reading of text; TBLCC = team-based learning comprehension check; TBLKA = team-based learning knowledge acquisition.

were implemented in the treatment classes, and very few elements of PACT were seen in the comparison classes.

Measures

Content knowledge, reading comprehension in the content area, and general reading comprehension were measured prior to and immediately following intervention. Follow-up data on retention of content knowledge were also collected 4 weeks and 12 weeks following intervention. Trained research personnel uninformed of participant conditions collected all data.

Assessment of Social Studies Knowledge (ASK). The ASK (Vaughn et al., 2013) assessment is a researcher-developed measure that includes two subtests. The first subtest, knowledge acquisition, is a 42-item, four-option, untimed multiple-choice test that measures content knowledge in the three units that composed the intervention (Gilded Age, Imperialism and World War I, The Twenties). With permission, items with known difficulty parameters were collected from released state and advanced placement social studies tests from Texas, Massachusetts, and The College Board. Researcher-developed vocabulary items were also included in the item set. The ASK knowledge acquisition measure was administered at pretest, posttest, 4 weeks following intervention, and again 12 weeks following intervention.

The second subtest—reading comprehension—is a 21-item, four-option, untimed multiple-choice test that measures reading comprehension in the content area. The assessment consists of three reading passages (Lexile range = 1210–1260; word count range = 348–392), each of which is related to content covered in the three 10-day cycles. Students read each passage silently and immediately answer seven multiple-choice questions about the passage. Reading comprehension items were researcher developed and measured students' ability to identify main ideas, understand vocabulary in context, identify cause and effect, and summarize. The ASK reading comprehension measure was administered at pretest and posttest.

A set of items for the ASK were selected after a series of pilot tests to validate the provided difficulty parameters, refine the instructions for test administrators, and estimate the amount of time necessary for administration (though untimed, we were concerned with test fatigue). The final set of items were selected following a series of item-level confirmatory factor analyses to evaluate model fit and estimate item parameters (Vaughn et al., 2013; Wanzek et al., 2014). We also estimated Cronbach's alpha on the measure with the current sample. Alpha coefficients for the ASK knowledge acquisition and reading comprehension measures were .90 and .84, respectively.

Gates–MacGinitie Reading Comprehension Subtest—4th edition. The Gates–MacGinitie Reading Comprehension subtest (MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2006) is a group-administered, timed (35 min) assessment of reading comprehension. The assessment consists of expository and narrative passages ranging in length from three to 15 sentences. Students read each passage silently and answer three to six multiple-choice questions related to the most recently read passage. As the students progress through the assessment, items increase in difficulty. The Gates–MacGinitie reading comprehension measure was administered at pretest and posttest. Internal consistency reliability ranges from .91 to .93, and alternate form reliability is reported as .80 to .87.

Data Analysis

We fit multilevel models to estimate the effects of treatment on students' acquisition of content-area knowledge, on their content-area reading comprehension, and on their general reading comprehension. Students were nested within classes and classes within teachers; classes were randomized to condition within teachers, meaning that each teacher taught both treatment and comparison classes. Pretest scores were grand-mean centered and used as Level-1 covariates (Enders & Tofighi, 2007). Treatment effects were modeled at Level 2. We calculated effect sizes as Hedges's g using the regression coefficient corresponding to the intervention effect as the numerator and the pooled standard deviation at posttest as the denominator. Site differences were modeled at Level 3. The interaction of site and condition was evaluated using the product term for the relevant cross-level interaction. The corresponding equations are as follows:

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{PRETEST} - gm_{ijk}) + e_{ijk}, \quad (1)$$

where i is student, j is class, k is teacher;

Y_{ijk} is the reading outcome of student i in class j and teacher k ;

π_{0jk} is the overall adjusted mean reading outcome of class j in teacher k after controlling for differences in pretest status;

$\pi_{1jk}(\text{PRETEST} - gm_{ijk})$ is the pretest covariate (grand mean centered);

e_{ijk} is the variance between students over classes in the adjusted mean reading outcome.

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(\text{COND})_{jk} + \beta_{02k}(\text{SITE} * \text{COND})_{jk} + r_{0jk} \quad (2)$$

$$\pi_{1jk} = \beta_{10k}, \quad (3)$$

where β_{00k} is the adjusted mean for school k ;

$\beta_{01k}(\text{COND})_{jk}$ is the treatment effect;

$\beta_{02k}(\text{SITE} * \text{COND})_{jk}$ is the cross-level term indicating the site by treatment interaction;

β_{10k} is the pretest covariate held constant across classes;

r_{0jk} is the variance between classes within teachers in the adjusted reading outcome.

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{SITE})_k + u_{00k} \quad (4)$$

$$\beta_{01k} = \gamma_{010} \quad (5)$$

$$\beta_{02k} = \gamma_{020} \quad (6)$$

$$\beta_{10k} = \gamma_{100}, \quad (7)$$

where γ_{000} is overall adjusted mean reading outcome in control schools in Florida;

$\gamma_{001}(\text{SITE})_k$ is fixed site effect;

γ_{010} is treatment effect constant across all teachers;

γ_{020} is site effect constant across all teachers;

γ_{100} is pretest covariate constant across all teachers;

u_{00k} is variance between schools in the adjusted mean reading outcome.

Attrition. Tables 3 and 4 summarize the percentage of missing data at each time point. We used two-way analyses of variance to evaluate the differential and potentially biasing effects of attrition (Cook & Campbell, 1979). Dependent variables consisted of pretest values on the

Table 3. Imputed means, standard deviations, and ranges for ASK knowledge acquisition

	Pretest			Posttest			Follow-Up 1			Follow-Up 2		
	<i>M</i>	<i>SD</i>	% Missing	<i>M</i>	<i>SD</i>	% Missing	<i>M</i>	<i>SD</i>	% Missing	<i>M</i>	<i>SD</i>	% Missing
Intervention ^a	18.45	6.15	8.76	24.89	7.97	17.09	23.55	8.93	26.06	23.87	8.37	48.93
Comparison ^b	17.45	5.63	11.24	21.66	7.34	17.87	21.23	8.11	25.64	21.23	7.65	52.16

Note. ASK = Assessment of Social Studies Knowledge.

^a*n* = 468. ^b*n* = 347.

Table 4. Imputed means, standard deviations, and ranges for Gates–MacGinitie and Ask Reading Comprehension

Measures	Pretest			Posttest		
	<i>M</i>	<i>SD</i>	% Missing	<i>M</i>	<i>SD</i>	% Missing
ASK reading comprehension in social studies						
Intervention ^a	10.44	4.12	8.33	10.91	4.00	16.88
Comparison ^b	9.79	3.83	7.49	10.36	4.02	18.73
Gates–MacGinitie reading comprehension						
Intervention ^a	100.39	11.70	4.91	96.81	12.86	18.37
Comparison ^b	99.09	10.20	4.89	95.33	12.33	19.56

Note. ASK = Assessment of Social Studies Knowledge.

^a*n* = 468. ^b*n* = 347.

primary reading variables (ASK knowledge acquisition, ASK reading comprehension, and Gates–MacGinitie reading comprehension). Independent variables in the analysis included treatment condition, completer status at different points in time (e.g., posttest, first follow-up, and second follow-up when appropriate) and the interaction of condition and completer status at different points in time. In this scenario, a significant main effect for the group of completers would indicate significant overall attrition and would represent a potential threat to external validity. A significant interaction would signify systematic condition-related differences across students who remained in the study and would be a potential threat to internal validity. On ASK knowledge acquisition, there were significant main effects for completer status at posttest and first follow-up ($p = .003$) but not at the second follow up ($p = .09$). There were no significant Condition \times Completer Status interaction effects at any of the time points (p values from .27 to .81). These findings suggest that although lower performing participants were more likely than higher performing students not to be present at posttest and at the first follow-up, there were no differences across treatment and comparison conditions in rates of attrition among the lower performing students. On ASK reading comprehension and Gates reading comprehension, there were no significant main effects for completer status ($p = .07$ and $p = .13$) and no significant Condition \times Completer Status interaction effect ($p = .38$ and $p = .80$).

To account for missing data, we used the multiple imputation algorithm provided in *Mplus* 7.1 to create 50 complete data sets. Multiple imputation uses all available information for each case, including demographic data, to estimate plausible values for missing data.

Growth Mixture Modeling. To examine patterns of response to treatment, we fit growth mixture models (Muthén et al., 2002; Nylund, Asparouhov, & Muthén, 2007) to represent population heterogeneity. We classified students into groups based on the combination of pretest scores *and* change in performance over time. Growth mixture modeling identifies unobserved subpopulations within a given sample based on qualitative differences in growth trajectories (e.g., Muthén, 2001). This contrasts with conventional latent growth models, which assume that a given sample is from a single population.

We compared four models, differing in number of classes (2-class model through 5-class model), using multiple indices to identify the most informative and parsimonious model. Indices included the Akaike Information Criterion (AIC; Akaike, 1987), Bayes information criteria (BIC), the Adjusted Bayesian Information Criterion (ABIC;

Nylund et al., 2007), and the index for entropy (Celeux & Soromenho, 1996). Models with the lowest AIC, BIC, and ABIC values indicate better fit to the data and increased probability of replication (Muthén, 2004). Latent class solutions were also compared in terms of their substantive interpretability.

RESULTS

Main Effects of Intervention on Student Achievement

We evaluated condition-related comparability at Time 1 using three-level analysis of pretest scores. On average, the treatment group scored higher at pretest on the measures of interest, including ASK knowledge acquisition ($\gamma_{01} = .96$, $SE = .40$, $p = .02$), ASK reading comprehension ($\gamma_{010} = .62$, $SE = .31$, $p = .04$), and Gates–MacGinitie reading comprehension ($\gamma_{01} = 1.78$, $SE = .71$, $p = .04$). Accordingly, pretest scores were included as Level-1 covariates in subsequent models. Treatment-specific means and standard deviations are presented in Tables 3 and 4. Fixed and random effects are in Table 5.

Content Knowledge. At posttest, students in the treatment group significantly outperformed students in the comparison group on the ASK knowledge acquisition ($\gamma_{010} = 2.80$, $SE = .67$, $p = .00$). This represents an effect size of .36 (Hedges's g). We also tested the possibility that treatment effects differed by teacher. Using a likelihood ratio test, we compared the model with treatment as a fixed factor to the model where treatment was estimated as a random effect. The deviance difference with 1 df was .09 ($p = .34$), suggesting that the addition of random effects did not improve model fit. The main effect of site was significant ($\gamma_{001} = 2.17$, $SE = .94$, $p = .02$), indicating that students in Site 1 had significantly higher posttest scores than students in Site 2. However, treatment effects did not interact with site ($\gamma_{001} = -.73$, $SE = .88$, $p = .41$), suggesting that students in both sites benefitted comparably from participation in reading intervention. At the first follow-up, the difference between the treatment and comparison groups was 1.89 ($p = .01$) and the effect size was .22. At the second follow-up, groups differed by 1.97 ($p = .01$). The effect size was .24.

Results from the unconditional models are presented in Table 5. The random effects for the ASK knowledge acquisition at posttest were significantly different from 0 at Level 1 and Level 2 of the model. More specifically, 9.87% of the variance in scores was attributable to classroom-level differences; 16% of the variance in scores was attributable to teacher-level differences, and the remaining 74.12% was attributable to student-level differences. At the first follow-up only the variance at student level was significantly different from 0, whereas at the second follow-up variances at all three levels were significantly different from 0.

Reading Comprehension. There were no significant differences on the ASK reading comprehension ($\gamma_{001} = .08$, $SE = .33$, $p = .81$; $g = .02$) or on the Gates–MacGinitie Reading Comprehension subtest ($\gamma_{001} = .48$, $SE = .99$, $p = .63$; $g = .04$). On the ASK reading comprehension, variances at the class (5.22%) and student level (83.61%) differed significantly from 0. Variance at the teacher level (11.17%) was not significantly different from 0. On the Gates–MacGinitie, variances at the teacher and student level differed significantly from 0, with 13.7% of the total variance at the teacher level and 82.76% at the student level. Variance at the class level (3.54%) did not differ significantly.

Table 5. Fixed and random effects by outcome

		Unconditional Model		Conditional Model		
Measures	Predictor	Estimate (<i>SE</i>)	<i>p</i>	Estimate (<i>SE</i>)	<i>p</i>	Effect Size
Fixed effects						
ASK knowledge acquisition at posttest	Intercept	23.42 (.98)	.00	21.11 (0.84)	.00	.36
	Pretest			.84 (0.04)	.00	
	Intervention			2.80 (0.67)	.00	
	Site			2.17 (0.94)	.02	
	Intervention × Site			−.73 (0.88)	.41	
ASK knowledge acquisition at Follow-Up 1	Intercept	22.46 (1.04)	.00	20.43 (0.79)	.00	.22
	Pretest			.97 (0.05)	.00	
	Intervention			1.89 (0.75)	.01	
	Site			3.02 (1.10)	.01	
	Intervention × Site			−1.28 (1.12)	.25	
ASK knowledge acquisition at Follow-Up 2	Intercept	22.79 (1.07)	.00	19.92 (0.83)	.00	.24
	Pretest			.83 (0.04)	.00	
	Intervention			1.97 (0.77)	.01	
	Site			4.54 (0.95)	.00	
	Intervention × Site			−.76 (0.96)	.43	
ASK reading comprehension	Intercept	10.63 (.42)	.00	10.46 (0.35)	.00	.02
	Pretest			.61 (0.04)	.00	
	Intervention			.08 (0.33)	.81	
	Site			.16 (0.50)	.74	
	Intervention × Site			.30 (0.43)	.48	
Gates-MacGinitie reading comprehension	Intercept	96.00 (1.37)	.00	96.39 (1.18)	.00	.04
	Pretest			.75 (0.03)	.00	
	Intervention			.48 (0.99)	.63	
	Site			−1.81 (1.36)	.18	
	Intervention × Site			.54 (1.40)	.70	

(Continued on next page)

Table 5. Fixed and random effects by outcome (*Continued*)

		Unconditional Model			Conditional Model		
		Variance (<i>SE</i>)	<i>p</i>	%	Variance (<i>SE</i>)	<i>p</i>	%
Random effects							
ASK knowledge acquisition at posttest	Level 1	46.57 (3.17)	.00	74.12	26.14 (1.73)	.00	86.96
	Level 2	6.20 (2.60)	.02	9.87	.48 (0.43)	.27	1.60
	Level 3	10.07 (7.09)	.16	16.01	3.44 (2.81)	.22	11.44
ASK knowledge acquisition at Follow-Up 1	Level 1	59.20 (4.51)	.00	77.77	31.46 (2.03)	.00	88.59
	Level 2	5.04 (2.82)	.07	6.61	1.06 (0.79)	.18	2.99
	Level 3	11.89 (7.55)	.12	15.62	2.99 (2.30)	.19	8.42
ASK knowledge acquisition at Follow-Up 2	Level 1	50.82 (3.20)	.00	74.53	30.43 (2.16)	.00	88.13
	Level 2	4.24 (2.08)	.04	6.12	.73 (0.71)	.30	2.11
	Level 3	13.16 (6.76)	.05	19.30	3.37 (2.02)	.10	9.76
ASK reading comprehension	Level 1	13.60 (.92)	.00	83.61	8.97 (0.47)	.00	94.30
	Level 2	0.85 (.41)	.04	5.22	.05 (0.13)	.73	.50
	Level 3	1.82 (.98)	.06	11.17	.49 (0.21)	.02	5.20
Gates–MacGinitie reading comprehension	Level 1	132.05 (7.01)	.00	82.76	75.74 (4.29)	.00	94.12
	Level 2	5.65 (5.09)	.27	3.54	1.41 (1.40)	.31	1.70
	Level 3	21.86 (10.12)	.03	13.70	3.32 (2.21)	.13	4.10

Note. ASK = Assessment of Social Studies Knowledge.

Differences in Treatment Effect for ASK Knowledge Acquisition

The fit indices, summarized in Table 6, suggested a two-group model (e.g., values for the AIC, BIC, and ABIC were the lowest for the two-group solution): (a) low baseline and no statistically significant growth over time (Group 1) and (b) high baseline in content knowledge and significant growth over time (Group 2). Group 1 (low performing) consisted of 481 students. The average pretest score for the ASK knowledge acquisition was 15.06 items correct, on average (i.e., independent of treatment condition), and the average growth from pretest to posttest was 2.61 ($p = .80$) items correct. Group 2 (high performing) had 334 students. The model-estimated mean on the ASK knowledge acquisition pretest was 22.27, and the predicted growth rate was 6.5 ($p < .001$), on average, for all students in this group, independent of condition. The treatment effect within the high-performing group (see Figure 1) was statistically significant as well. Growth in the treatment group was greater by 2.6 ($\beta = 2.6$, $p = .00$) raw score points in the treatment group relative to the comparison. There were no significant treatment effects in the low-performing group ($\beta = 1.9354$, $p = .64$).

Table 6. Fit indices from model testing

Classes	AIC	BIC	ABIC	Entropy	Class Composition
2	10224	10304	10250	.60	$n_1 = 59\%, n_2 = 41\%$
3	10218	10340	10258	.69	$n_1 = 1.3\%, n_2 = 40.8\%, n_3 = 57.9\%$
4	10219	10384	10273	.74	$n_1 = 2.1\%, n_2 = 42.8\%, n_3 = 53.8\%, n_4 = 1.3\%$
5	10220	10427	10288	.71	$n_1 = 1.94\%, n_2 = 16.06\%, n_3 = 40.45\%,$ $n_4 = 38.15\%, n_5 = 3.4\%$

Note. The percentages reported in the far right column represent the distribution of the participants across the classes for that particular model. AIC = Akaike Information Criteria; BIC = Bayesian Information Criterion; ABIC = Sample size adjusted Bayesian Information Criterion.

DISCUSSION

The purpose of this study was to examine the effects of the PACT intervention implemented in 11th-grade social studies classes. Overall, following three instructional units in U.S. History classes, students in the PACT intervention significantly improved their content knowledge. Students in classes receiving the PACT intervention performed approximately one third of a standard deviation higher than students receiving typical social studies instruction in the same content. Of importance, the treatment students maintained these gains in content knowledge 4 and 12 weeks following the intervention. The long-term effects of the PACT intervention demonstrate the stability of the knowledge acquisition for students receiving the treatment. The goal of any instruction is to improve student outcomes that are maintained over time. Our findings provide support for the PACT intervention as a way to substantially improve students’ content knowledge acquisition and retention in high school history classes.

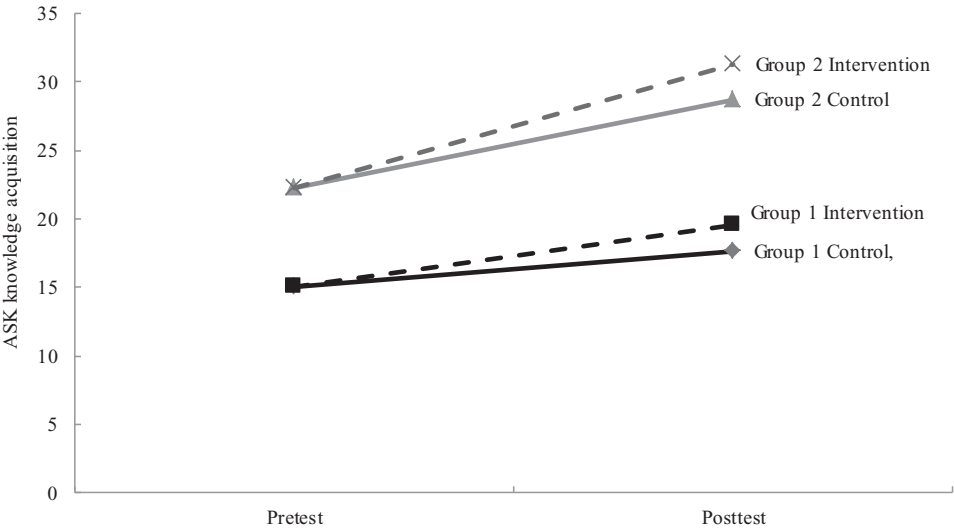


Figure 1. Estimated pretest and posttest mean scores on Assessment of Social Studies Knowledge (ASK) for the intervention and comparison groups from the two-class model.

The PACT intervention did not improve students' reading comprehension above the typical instruction provided. We hypothesized that the intervention's focus on text in the social studies domain would improve students' comprehension of novel social studies text. However, the findings did not support this hypothesis. Because of the focus of the intervention on knowledge acquisition and content understanding, we hypothesized the intervention would not significantly affect students' general reading comprehension outside of the domain. We found there were no differences between the study conditions on the standardized measure of reading comprehension. Given that students in the treatment did not significantly outperform students in the comparison condition on comprehension of social studies text, it is not surprising there were also no differences in general reading comprehension. Additional text focus and instruction may be needed in both content area literacy as well as general reading to improve students' reading comprehension.

The effect size for knowledge acquisition of .36 in this study was larger than the effect size of .19 noted in the study of the TBL component alone (Wanzek et al., 2014). Although not a direct comparison, these findings suggest the magnitude of effects for the PACT intervention was larger than when only TBL was implemented in 11th-grade social studies classes. Thus, although the literacy components of vocabulary, text reading, and use of text for evidence did not improve students' reading comprehension beyond typical social studies practice, these components may have been necessary to reach the level of knowledge acquisition that was observed.

We also found differential student response to instruction in knowledge acquisition. Students with higher pretest scores entering unit implementation made the greatest gains during the unit instruction. In addition, this group of students significantly benefitted from the PACT implementation above and beyond the growth seen in typical instruction. Students with lower pretest scores made fewer gains in content knowledge during instruction. These findings suggest that students with some background knowledge prior to unit implementation benefitted the most from the instruction. Students with less background knowledge may need additional supports to better access the content or instruction adequately to make sufficient gains in content knowledge. Knowledge acquisition was statistically similar in the PACT treatment and typical practice for this lower group, though there was a more positive trend for the students who received the PACT treatment. Wanzek et al. (2014) reported a similar finding for the effects of TBL implementation with 11th graders. Students with higher pretest scores prior to the unit instruction made more gains during the unit, particularly when they received the TBL treatment. More intensive attention to students' background knowledge and the supports these students require to fully engage in the unit instruction may be needed.

Implications and Future Research

The findings of this study suggest the potential of the PACT intervention to increase content knowledge in the social studies for high school students. The consistent findings for improved content knowledge from this study as well as the two studies examining PACT implementation in eighth grade (Vaughn et al., 2013; Vaughn et al., in press) provide an evidence-base for PACT as a useful practice for teachers to consider within their social studies content delivery. This study occurred over a relatively short period (30 days of instruction) and would benefit from future research examining impacts from longer implementations.

The differential results for students with the lowest pretest scores, a finding also noted by Wanzek et al. (2014), indicates a need for further research in general education high school content area instruction. Neither current typical practices nor the PACT intervention sufficiently met the needs of students with low pretest scores. Future research could examine the supports these students require in their general education to make more substantial gains in their learning of social studies content and better prepare them for college or career engagement.

The teachers in this study received both training and in-class support in implementing the components of the PACT intervention. The more student-centered approach to instruction seen in the PACT components can be challenging for teachers to successfully incorporate in K-12 classrooms (Beck & Eno, 2012). Additional research on assisting teachers in gaining facility with and sustaining this type of instruction is warranted given the consistent findings of more traditional types of instruction (e.g., lecture, note taking) that have persisted in the social studies (Bolinger & Warren, 2007; Chiodo & Byford, 2004; Paxton, 1999; Swanson et al., in press).

Overall, the findings of this study provide evidence that student gains can be realized when these more student-centered practices are implemented. In addition, these practices provide teachers with a way to infuse the social studies with reading practices that enhance content acquisition.

FUNDING

The research reported here was supported by Grant R305F100013 from the Institute of Education Sciences, U. S. Department of Education. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

REFERENCES

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317–332. doi:10.1007/bf02294359
- Barton, K. C., & Levstik, L. S. (2003). Why don't more history teachers engage students in interpretation? *Social Education*, 67, 358–361.
- Beck, D., & Eno, J. (2012). Signature pedagogy: A literature review of social studies and technology research. *Computers in the Schools*, 29, 70–94. doi:10.1080/07380569.2012.658347
- Berkeley, S., Marshak, L., Mastropieri, M. A., & Scruggs, T. E. (2011). Improving student comprehension of social studies text: A self-questioning strategy for inclusive middle school classes. *Remedial and Special Education*, 32, 105–113. doi:10.177/0741932510361261
- Bolinger, K., & Warren, W. J. (2007). Methods practiced in social studies instruction: A review of public school teachers' strategies. *International Journal of Social Education*, 22, 68–84.
- Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13, 195–212. doi:10.1007/BF01246098
- Chiodo, J. J., & Byford, J. (2004). Do they really dislike social studies? A study of middle school and high school students. *Journal of Social Studies Research*, 28, 16–26.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32, 9–13. doi:10.3102/0013189x032001009
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago, IL: Rand McNally.
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12, 121–138. doi:10.1037/1082-989x.12.2.121

- Gwet, K. (2001). *Handbook of inter-rater reliability: How to estimate the level of agreement between two or multiple raters*. Gaithersburg, MD: STATAXIS.
- Hall, L. A. (2005). Teachers and content area reading: Attitudes, beliefs, and change. *Teaching and Teacher Education*, 21, 403–414.
- Hein, V., Smerdon, B., Lebow, M., & Agus, J. (2012). *Goals and expectations for college and career readiness: What should students know and be able to do?* Retrieved from http://www.betterhighschools.org/CCR/documents/NHSCCCRGGoalsExpectationsBrief_2012.pdf
- Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas*. Washington, DC: Alliance for Excellent Education.
- Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, 2, 88–110. doi:10.1080/19345740802539325
- Lee, C. D., & Spratley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy*. New York, NY: Carnegie Corporation of New York.
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2006). *Gates–MacGinitie reading tests (4th ed.)*. Rolling Meadows, IL: Riverside.
- Michaelsen, L. K., & Sweet, M. (2011). Team-based learning. *New Directions for Teaching and Learning*, 128, 41–51. doi:10.1002/tl.467
- Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In L. M. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 291–322). Washington, DC: American Psychological Association. doi:10.1037/10409-010
- Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 345–368). Thousand Oaks, CA: Sage. doi:10.4135/9781412986311.n19
- Muthén, B. O., Brown, C. H., Masyn, K., Jo, B., Khoo, S., Yang C., . . . Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics*, 3, 459–475. doi:10.1093/biostatistics/3.4.459
- National Council for the Social Studies. (2008). *Curriculum guidelines for social studies teaching and learning: A position statement of the National Council for the Social Studies*. Retrieved from http://www.socialstudies.org/system/files/files/Curriculum_Guidelines_SocialStudies_Teaching_and_Learning.pdf
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards: Social studies*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.
- New York City Department of Education. (2013, December 10). *Planning for success: Supporting transitions through high school to college and career*. Retrieved from http://schools.nyc.gov/NR/ronlyres/9FDD0841-D54D-4D16-ADAA-9F116B79F645/0/ParentGuide_forCCL_updated080712.pdf
- Nylund, K. L., Asparouhov, T. & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14, 535–569. doi:10.1080/10705510701575396
- O'Brien, D. G., Moje, E. B., & Stewart, R. (2001). Exploring the context of secondary literacy: Literacy in people's everyday school lives. In E. B. Moje & D. G. O'Brien (Eds.), *Construction of literacy: Studies of teaching and learning in and out of secondary schools* (pp. 105–24). Mahwah, NJ: Erlbaum.
- O'Brien, D. G., Stewart, R. A., & Moje, E. B. (1995). Why content literacy is difficult to infuse into the secondary school: Complexities of curriculum, pedagogy, and school culture. *Reading Research Quarterly*, 30, 442–463. doi:10.2307/747625
- Paxton, R. J. (1999). A deafening silence: History textbooks and the students who read them. *Review of Educational Research*, 69, 315–339. doi:10.3102/00346543069003315

- Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, 78, 40–59.
- Stahl, S. A., Hynd, C. R., Glynn, S. M., & Carr, M. (1996). Beyond reading to learn: developing content and disciplinary knowledge through texts. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. 139–164). Mahwah, NJ: Erlbaum.
- Swanson, E. A., Wanzek, J., McCulley, L., Stillman, S., Vaughn, S., Simmons, D., . . . Hairrell, A. (in press). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading and Writing Quarterly*.
- Vaughn, S., & Fletcher, J. M. (2012). Response to intervention with secondary school students with reading difficulties. *Journal of Learning Disabilities*, 45, 244–256. doi:10.1177/0022219412442157
- Vaughn, S., Roberts, G., Swanson, E. A., Wanzek, J., Fall, A., & Stillman-Spisak, S. J. (in press). Improving middle school students' knowledge and comprehension in social studies: A replication. *Educational Psychology Review*.
- Vaughn, S., Swanson, E., Roberts, G., Wanzek, J., Stillman-Spisak, S. J., Solis, M., & Simmons, D. (2013). Improving reading comprehension and social studies knowledge in middle school. *Reading Research Quarterly*, 48, 77–93. doi:10.1002/rrq.039
- Wanzek, J., Vaughn, S., Kent, S. C., Swanson, E. A., Roberts, G., Haynes, M., . . . Solis, M. (2014). The effects of team-based learning on social studies knowledge acquisition in high school. *Journal of Research on Educational Effectiveness*, 7, 183–204. doi:10.1080/19345747.2013.836765